

A note on constructive treatment of eigenvectors

P. Osinenko*, G. Devadze, and S. Streif

Laboratory for Automatic Control and Systems Dynamics
Technische Universität Chemnitz

* pavel.osinenko@etit.tu-chemnitz.de

Abstract

The eigenvalue problem plays a central role in linear algebra and its applications in control and optimization methods. In particular, many matrix decompositions rely upon computation of eigenvalue-eigenvector pairs, such as diagonal or Jordan normal forms. Unfortunately, numerical algorithms computing eigenvectors are prone to errors. Due to uncomputability of eigenpairs, perturbation theory and various regularization techniques only help if the matrix at hand possesses certain properties such as the absence of non-zero singular values, or the presence of a distinguishable gap between the large and small singular values. Posing such a requirement might be restrictive in some practical applications. In this note, we propose an alternative treatment of eigenvectors which is approximate and constructive. In comparison to classical eigenvectors whose computation is often prone to numerical instability, a constructive treatment allows addressing the computational uncertainty in a controlled way.

Index terms— Eigenvalues, eigenvectors, constructive analysis, approximate solutions, fundamental theorem of algebra

1 Introduction

Let \mathbf{A} be a complex-valued $n \times n$ matrix. Its characteristic polynomial is given as:

$$P_A(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}). \quad (1)$$

An eigenpair $(\mathbf{v}_j, \lambda_j)$ is a root $\lambda_j, j \in \{1, \dots, n\}$ of (1) and a vector \mathbf{v}_j satisfying:

$$\mathbf{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j. \quad (2)$$

Computing eigenpairs is crucial for many control and optimization methods because these methods are often based on certain matrix decompositions such as diagonal normal form, Jordan form or singular value decompositions (SVD). For

example, SVD is used in H_2 , H_∞ and μ -optimal controllers by decoupling the original system into lower-dimensional ones, e.g. [1], and requires computing left and right eigenvectors. SVD is also used in deriving reduced models and controllers which preserve important system properties such as closed-loop stability, observability or controllability [2, 3]. In model predictive control, SVD of the cost function's Hessian is often used (see, e.g. [4]). System identification is another field of application of SVD. For instance, Zhang et al. [5] used SVD on extended Kalman filter to cope with numerical ill-conditioning. Eigenvectors are used in computing Jordan normal form for stability analysis of linear systems [6]. For a survey on controller design based on matrix normal forms refer to [7].

Computing eigenvectors amounts to finding non-trivial solutions to the system of linear equations (2). The classical method of solving systems of linear equations is Gaussian elimination (see, e.g., [8]). In general, solving a system of linear equations $\mathbf{A}\mathbf{v} = \mathbf{b}$, where \mathbf{A} is the coefficient matrix, \mathbf{b} is the data vector, and \mathbf{v} is an unknown vector, amounts to comparing real numbers to zero. In computations, real numbers are represented by computer programs or algorithms that compute the respective approximations. There is, however, no algorithm that can decide whether $\alpha = \beta$ or $\alpha \neq \beta$ for arbitrary real numbers α and β . Such an algorithm would be equivalent to solving the problem of deciding whether an arbitrary computer program terminates or not – which is impossible as shown by Turing [9]. When performing elementary matrix row operations, one might have to deal with numbers that are small relative to the machine precision without actually knowing whether they are actually zero or not. In practice, small numbers are either set to zero or to machine epsilon. However, this may lead to a large deviation of an approximate solution from the true one [10].

Besides the problem of approximating small numbers when solving $\mathbf{A}\mathbf{v} = \mathbf{b}$, a typical problem is that the coefficient matrix \mathbf{A} is singular. A comprehensive analysis of these ill-posed problems was carried out within perturbation theory [11, 12, 10, 13, 14]. The techniques that cope with solving such ill-posed problems approximately, mainly use regularization or truncated singular value decomposition (TSVD). In brief, if the coefficient matrix \mathbf{A} of a problem $\mathbf{A}\mathbf{v} = \mathbf{b}$ is close to singular, it is substituted by some other matrix which is regular. For instance, Tikhonov regularization [12] suggests to use $(\mathbf{A}\mathbf{A}^* + h^2\mathbf{I})^{-1}\mathbf{A}^*\mathbf{b}$ as an approximate solution to $\mathbf{A}\mathbf{v} = \mathbf{b}$. In TSVD, \mathbf{A} is decomposed into $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$, where \mathbf{U} and \mathbf{V}^* are unitary matrices, and $\mathbf{\Sigma}$ is a diagonal matrix of singular values $\{\sigma_j\}_j$ of \mathbf{A} , i.e., square roots of the eigenvalues of $\mathbf{A}^*\mathbf{A}$. The singular values, which are small relative to the given threshold, are set to zero (“truncated”). The resulting approximate solution to $\mathbf{A}\mathbf{v} = \mathbf{b}$ may be found as follows:

$$\hat{\mathbf{v}} = \sum_{j \in I_1} \frac{\mathbf{u}_j^T (\mathbf{A}^T \mathbf{b})}{\sigma_j} \mathbf{v}_j + \sum_{j \in I_2} t_j \mathbf{v}_j, \quad (3)$$

where $\mathbf{u}_j, \mathbf{v}_j$ are the j -th column of \mathbf{U} and \mathbf{V} , respectively, with index sets $I_1 := \{j : \sigma_j \neq 0\}$ and $I_2 := \{j : \sigma_j = 0\}$, and t_j are free variables associated

with those singular values that have been truncated.

In practice, such regularization techniques work well only when there is a distinguishable gap between the singular values that are large (i.e. contained in I_1) and those which are near zero (i.e. contained in I_2) relative to the machine precision. In general, any approximate solution of this kind may yield large approximation errors. The bound on a solution is proportional either to the so-called condition number $\|\mathbf{A}^{-1}\|\|\mathbf{A}\|$, or, equivalently, to the reciprocal of the smallest singular value of \mathbf{A} , or to the said gap between singular values. Here $\|\mathbf{A}\| \triangleq \sup\{\|\mathbf{A}\mathbf{v}\| : \mathbf{v} \in \mathbb{C}^n \wedge \|\mathbf{v}\| = 1\}$ denotes the operator norm of \mathbf{A} and $\|\cdot\|$ is the Euclidean vector norm. Then, if \mathbf{A} has actually a zero singular value or if the said gap is close to the machine epsilon, regularization techniques may yield poor approximate solutions [10].

When looking numerically for non-trivial solutions to $\mathbf{A}\mathbf{v} = \mathbf{0}$, one often deals with some approximation $\hat{\mathbf{A}}$ to \mathbf{A} . Let $\hat{\mathbf{A}} = \mathbf{A} + \mathbf{E}$, where \mathbf{E} is some perturbation matrix. Eigenvectors of $\hat{\mathbf{A}}$ then satisfy:

$$\hat{\mathbf{v}}_j = \mathbf{v}_j + \sum_{\substack{k=1 \\ k \neq j}} \frac{\mathbf{v}_k^T \mathbf{E} \mathbf{v}_k}{\lambda_j - \lambda_k} \mathbf{v}_j, \quad (4)$$

where $\{\lambda_j\}_j$ and $\{\mathbf{v}_j\}_j$ are the original eigenvalues and eigenvectors. For details and numerical algorithms, refer to [15]. Some other algorithms, such as approximation of eigenvectors by least squares [16], rely on the knowledge of the first non-zero singular value of \mathbf{A} . Therefore, solving systems of linear equations $\mathbf{A}\mathbf{v} = \mathbf{b}$, and finding eigenvectors in particular, requires certain properties of the original problem to be satisfied which might be restrictive in some applications. To approach the problem from a computational point of view, linear algebra was also addressed in computable analysis as developed by Weihrauch [17]. Computable analysis investigates mathematical objects that can be computed in principle, i.e., by an algorithm that is guaranteed to yield a correct answer. Brattka et al. [18] investigated the problem of computability of solutions to a problem $\mathbf{A}\mathbf{v} = \mathbf{b}$. They concluded that the solutions are computable provided that the rank of \mathbf{A} is known in advance. In general, however, the rank is uncomputable [19]. There also exist constructive frameworks of algebra. For a comprehensive course, please refer to a canonical book on the subject by Mines et al. [20]. The eigenvalue problem was investigated in constructive mathematics for particular linear operator decompositions. For example, Ye [21] investigated the spectral decomposition using constructive functional analysis. He used an assumption that all the eigenvalues be given with their multiplicities known beforehand. With this assumption, he was able to exactly reconstruct the spectral theorem.

In this note, we are neither concerned with *exact* eigenvectors, nor do we seek for new bounds on approximate solutions in the spirit of perturbation theory. Instead, we suggest to treat eigenvectors from the standpoint of constructive mathematics, as developed by Bishop et al. [22]. The advantage of this approach is that the derived proofs may be translated into effective algorithms

that always output a correct result [22]. This, in our view, is especially important in applications such as control engineering. We do not look for solutions to (2), but propose to use the following notion of an approximate eigenvector:

$$\|\mathbf{A}\hat{\mathbf{v}} - \hat{\lambda}_j\hat{\mathbf{v}}\| \leq \varepsilon, \quad (5)$$

where ε is the predefined accuracy, $\hat{\mathbf{v}}$ is an ε -eigenvector with $\|\hat{\mathbf{v}}\| = 1$, and $\hat{\lambda}_j$ is an approximate eigenvalue of \mathbf{A} in the sense that $\det(\mathbf{A} - \hat{\lambda}_j\mathbf{I}) \leq \varepsilon$. We constructively prove existence of ε -eigenvectors using the algebraic closure of rational complex numbers and continuity of roots of polynomials. The results derived in the present note do not require a prior knowledge of constructive functional analysis or measure theory. First, we discuss certain important notions of the constructive analysis in the preliminaries section, then we constructively verify that roots of polynomials are continuous in coefficients, and finally we prove existence of ε -eigenvectors. We conclude by pointing out the importance and usages of ε -eigenvectors in control and optimization.

2 Preliminaries

The aim of this section is to briefly introduce certain key notions of constructive mathematics that will be needed in this note. For a comprehensive description, refer to [22, 23, 24]. The central notion is an **operation** which is an algorithm that produces a unique result in a finite number of steps for each input from its domain. For example, a **real** number x is a *regular* Cauchy sequence of rational numbers in the sense that

$$\forall n, m \in \mathbb{N} \left[|x(n) - x(m)| \leq \frac{1}{n} + \frac{1}{m} \right]$$

where $x(n)$ is an *operation* that produces the n -th rational approximation to x . A **set** is a pair of operations: \in determines that a given object is a member of the set, and $=$ determines whether two given set members are equal. Existence and universal quantifiers are interpreted as follows: $\exists x \in A [\varphi(x)]$ means that an operation was derived that constructs an instance x along with a proof of $x \in A$ and a proof of the logical formula $\varphi(x)$ as *witnesses*; $\forall x \in A [\varphi(x)]$ means that an operation was derived that proves $\varphi(x)$ for any x provided with a witness for $x \in A$. A **multiset** is a finite collection of objects that allows repeated elements. We denote multisets by $\{\cdot\}$. For example $\{1, 1\}$ is a multiset. A function $f : [a, b] \rightarrow \mathbb{R}$ is **uniformly continuous** if there exists an operation $\omega : \mathbb{Q} \rightarrow \mathbb{Q}$ called **modulus of continuity** such that $\forall \varepsilon \in \mathbb{Q} \forall x, y \in [a, b] [|x - y| \leq \omega(\varepsilon) \implies |f(x) - f(y)| \leq \varepsilon]$.

A **complex number** $x + yi$ is identified with a pair (x, y) of real numbers. The set of complex numbers \mathbb{C} is a metric space with the metric $\rho(z, w) \triangleq |z - w|$. A **sphere** centered at z with the radius $r > 0$ is denoted by $\mathcal{S}(z, r)$. Let K respectively U be a compact respectively any set in \mathcal{C} respectively, and let $r > 0$. We define

$$K_r \triangleq \{z \in \mathbb{C} : \rho(z, K) \leq r\}$$

The set K is well-contained in U if there exists $r > 0$ such that $K_r \subseteq U$. A **path** γ is a uniformly continuous function $\gamma : [a, b] \rightarrow \mathbb{C}$ such that:

$$\exists t_0, \dots, t_n, n > 0 [t_0 = a \leq t_1 \leq \dots \leq t_n \leq b],$$

and γ is differentiable on each interval $[t_i, t_{i+1}]$, $i = 0, \dots, n-1$. A path is closed if $\gamma(a) = \gamma(b)$. By $\text{car}(\gamma)$, we denote the closure of the range of γ . Notice that $\text{car}(\gamma)$ is compact and $\rho(z, \text{car}(\gamma))$ exists for any z since γ is uniformly continuous. If a function f is differentiable on $\text{car}(\gamma)$ and γ is closed, then

$$\int_{\gamma} f' dz = 0.$$

If γ is a closed path, $z_0 \in \mathbb{C}$, and $\rho(z_0, \text{car}(\gamma)) > 0$ then:

$$\int_{\gamma} \frac{dz}{z - z_0} = 2\pi i \mathcal{W}(\gamma, z_0).$$

for some integer $\mathcal{W}(\gamma, z_0)$ called the **winding number** of γ around z_0 . We will denote the number of zeros of a function f within a closed path γ by $\mathcal{N}(f, \gamma)$. The **length** of a path γ is denoted by $\text{len } \gamma$ and defined by $\int_{\gamma} 1 dz$. A function $P(z) = \sum_{k=0}^n a_k z^k$ is a **polynomial** of degree n if $a_n \neq 0$. It is **monic** if $a_n = 1$. A continuous function f on an open set U is **analytic** on U , if $\int_{\gamma} f dz = 0$ for any triangular path γ going through some points z_1, z_2, z_3 such that the convex hull of z_1, z_2, z_3 is a subset of U . In the next section, we present necessary lemmas, required for the main theorem.

3 Technical lemmas

In this section, we show that continuity of roots of polynomials admits a constructive proof. Most of the classical derivations are easily transferable into the constructive setup.

Let $P(z)$ be a monic polynomial of degree n over \mathbb{C} . By the constructive Fundamental Theorem of Algebra (FTA) [22, p. 156], it has n roots. Therefore,

$$P(z) = z^n + \sum_{k=0}^{n-1} a_k z^k = \prod_{k=1}^n (z - z_k)$$

Let $\hat{P}(z)$ be some monic polynomial of degree n whose coefficients \hat{a}_k approximate those of $P(z)$ in the sense that $\forall k [|a_k - \hat{a}_k| \leq \varepsilon]$. The polynomial $\hat{P}(z)$ may be decomposed as well:

$$\hat{P}(z) = z^n + \sum_{k=0}^{n-1} \hat{a}_k z^k = \prod_{k=1}^n (z - \hat{z}_k) \quad (6)$$

We look for a bound δ , that depends on ε such that $\forall k [|z_k - \hat{z}_k| \leq \delta(\varepsilon)]$. That is, if the coefficients of $\hat{P}(z)$ are close to the coefficients of $P(z)$, then the respective

roots are close as well. To show this, we require several technical lemmas. The first is:

Lemma 1. *Let $P(z)$ be a polynomial of degree n , and let $\bar{\mathcal{B}}$ be some closed ball well-contained in an open set $U \subset \mathbb{C}$ with the boundary \mathcal{S} . Let γ denote a closed path such that $\text{car}(\gamma) = \mathcal{S}$. Assume that $P(z)$ has no roots on $\text{car}(\gamma)$. It follows that:*

$$\frac{1}{2\pi i} \oint_{\gamma} \frac{P'(z)}{P(z)} dz = \mathcal{N}(P, \gamma)$$

Proof. By the FTA, $P(z)$ has a multiset of n roots $\{z_1, \dots, z_n\}$. Without loss of generality, we may assume that $P(z)$ is monic. Then, it can be expressed as $P(z) = (z - z_1)P_1(z)$ where $P_1(z) = \prod_{k=2}^n (z - z_k)$. Therefore, $P'(z) = P_1(z) + (z - z_1)P_1'(z)$. It follows that:

$$\begin{aligned} \frac{1}{2\pi i} \oint_{\gamma} \frac{P'(z)}{P(z)} dz &= \frac{1}{2\pi i} \oint_{\gamma} \left(\frac{1}{(z - z_1)} + \frac{P_1'(z)}{P_1(z)} \right) dz \\ &= \frac{1}{2\pi i} \oint_{\gamma} \frac{1}{(z - z_1)} dz + \frac{1}{2\pi i} \oint_{\gamma} \frac{P_1'(z)}{P_1(z)} dz. \end{aligned}$$

The first integral equals $\mathcal{W}(\gamma, z_1) = 1$ if z_1 is in $\bar{\mathcal{B}}$, and 0 otherwise. Now, proceeding by finite induction, we obtain the result. \square

The next lemma shows that a parametric integral over a path is uniformly continuous.

Lemma 2. *Let $f(t, z)$ be a uniformly continuous function on $[t_0, t_1] \times U$, where $t_0 < t_1, t_0, t_1 \in \mathbb{R}$, and U is an open set in \mathbb{C} . Let $\gamma : [0, 1] \rightarrow \mathbb{C}$ be a path in U . Define $I(t) := \int_{\gamma} f(t, z) dz$. It follows that $I(t)$ is a uniformly continuous function of t .*

Proof. Let $\tau_1, \tau_2, \tau_1 \leq \tau_2$ be two points in $[t_0, t_1]$. First, observe that:

$$|I(\tau_1) - I(\tau_2)| = \left| \int_{\gamma} (f(\tau_1, z) - f(\tau_2, z)) dz \right| \leq \int_{\gamma} |f(\tau_1, z) - f(\tau_2, z)| dz.$$

Let ω_f be the modulus of continuity of f . It follows that if $|\tau_1 - \tau_2| \leq \omega_f(\varepsilon)$ for some $\varepsilon > 0$, then $|f(\tau_1, z) - f(\tau_2, z)| \leq \varepsilon$. Therefore,

$$\int_{\gamma} |f(\tau_1, z) - f(\tau_2, z)| dz \leq \int_{\gamma} \varepsilon dz = \varepsilon \text{len}(\gamma).$$

The result follows by setting $\omega_I(\varepsilon) := \omega_f\left(\frac{\varepsilon}{\text{len}(\gamma)}\right)$. \square

We require a variant of the Rouché's Theorem [25, p. 131] for polynomials:

Lemma 3. *Let $P(z)$ and $Q(z)$ be polynomials of degree n , and let $\bar{\mathcal{B}}$ be some closed ball in \mathbb{C} with the boundary \mathcal{S} . Let γ denote a closed path such that $\text{car}(\gamma) = \mathcal{S}$. Assume that $P(z)$ has no roots on $\text{car}(\gamma)$. If $|Q(z)| < |P(z)|$ for all z in $\bar{\mathcal{B}}$, then $\mathcal{N}(P, \mathcal{S}) = \mathcal{N}(P + Q, \mathcal{S})$.*

Proof. Define the following function:

$$I(t) := \frac{1}{2\pi i} \oint_{\gamma} \frac{P'(z) + tQ'(z)}{P(z) + tQ(z)}, t \in [0, 1].$$

Since $P(z)$ has no roots on $\text{car}(\gamma)$, $P(z) + tQ(z)$ also has no roots on $\text{car}(\gamma)$. By Lemma 1, $I(t) = \mathcal{N}(P + tQ, \mathcal{S}) \in \mathbb{N}$. The condition $|Q(z)| < |P(z)|$ on $\bar{\mathcal{B}}$ implies that for any z in $\bar{\mathcal{B}}$, there exists a positive rational complex number p such that:

$$|Q(z)| < p < |P(z)|.$$

By the density of rational complex numbers in \mathbb{C} , there exists also a positive q such that the following inequality holds:

$$|Q(z)| < p < q < |P(z)|.$$

By the triangle inequality, we have:

$$|P(z) + tQ(z)| > |P(z)| - t|Q(z)| > |P(z)| - |Q(z)| > q - p.$$

It follows that $\frac{1}{|P(z) + tQ(z)|}$ is analytic on $\text{car}(\gamma)$. Hence, $\frac{P'(z) + tQ'(z)}{P(z) + tQ(z)}$ is analytic on $\text{car}(\gamma)$ as well. By Lemma 2, $I(t)$ is uniformly continuous. But since $I(t)$ is an integer for all t , $I(t)$ is a constant function. Therefore, $I(0) = I(1)$ which means that $\mathcal{N}(P, \mathcal{S}) = \mathcal{N}(P + Q, \mathcal{S})$. \square

We are ready to prove the following lemma:

Lemma 4. *Let $P(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a monic polynomial. For any $\varepsilon > 0$, there exists a monic polynomial $\hat{P}(z) = z^n + \sum_{k=0}^{n-1} \hat{a}_k z^k$ and $\delta > 0$, that depends on ε , such that $\forall k$ $[|z_k - \hat{z}_k| \leq \varepsilon]$, where z_k, \hat{z}_k are a root of $P(z)$ and a root of $\hat{P}(z)$ respectively, and $\forall k$ $[|a_k - \hat{a}_k| \leq \delta(\varepsilon)]$.*

Proof. By the FTA, $P(z)$ has a multiset of n roots $\{z_1, \dots, z_n\}$. By approximating z_1, \dots, z_n sufficiently, we can find some m disjoint closed balls $\{\bar{\mathcal{B}}(c_j, n\varepsilon)\}_j$ of radius $n\varepsilon$ each containing some m_j points of $\{z_1, \dots, z_n\}$, and such that no root is on the boundary of the respective ball. Let

$$s_j := \inf_{\mathcal{S}(c_j, n\varepsilon)} |P(z)|, j = 1, \dots, m.$$

Notice that infima and suprema of polynomials on $\mathcal{S}(c_j, n\varepsilon)$ exist since $\mathcal{S}(c_j, n\varepsilon)$ are compact sets. Further, let

$$p_j := \sup_{\mathcal{S}(c_j, n\varepsilon)} \left\{ 1 + \sum_{k=1}^{n-1} |z^{n-k}| \right\} = \sup_{\mathcal{S}(c_j, n\varepsilon)} \{1 + |z^{n-1}| + |z^{n-1}| + \dots + |z|\}.$$

Let δ be such that $\forall j = 1, \dots, m \left[\delta < \frac{s_j}{p_j} \right]$ and $\forall k = 1, \dots, n-1 \left[|a_k - \hat{a}_k| \leq \delta \right]$. Then, it follows that if $z \in \{\bar{\mathcal{B}}(c_j, n\varepsilon)\}_j$:

$$\begin{aligned} |\hat{P}(z) - P(z)| &= \left| \sum_{k=0}^{n-1} (\hat{a}_k - a_k) z^k \right| \leq \sum_{k=0}^{n-1} |\hat{a}_k - a_k| |z^k| \leq \\ &\delta \sum_{k=0}^{n-1} |z^k| \leq \delta p_j < s_j \leq |P(z)| \end{aligned}$$

Therefore, $\forall j = 1, \dots, m \left[z \in \bar{\mathcal{B}}(c_j, n\varepsilon) \implies |P(z)| > |\hat{P}(z) - P(z)| \right]$. By Lemma 3, $\mathcal{N}(P, \mathcal{S}(c_j, n\varepsilon)) = \mathcal{N}(\hat{P}, \mathcal{S}(c_j, n\varepsilon)) = m_j$ since $(\hat{P}(z) - P(z)) + P(z) = \hat{P}(z)$. It follows that for a given ε , there is a δ that establishes the desired property. Notice that by the density of rational complex numbers, there exists such a polynomial $\hat{P}(z)$ with rational coefficients. \square

Using the lemmas in this section, and algebraic closure of rational complex numbers, we are ready to prove the existence of approximate eigenvectors.

4 Constructive ε -eigenvectors and discussion

Theorem 1. *Let A be a complex-valued $n \times n$ matrix with the characteristic polynomial $P_A(\lambda)$. For any $\varepsilon > 0$, there exist some m linearly independent vectors $\hat{\mathbf{v}}_1, \dots, \hat{\mathbf{v}}_m$ and complex numbers $\hat{\lambda}_1, \dots, \hat{\lambda}_m$ such that:*

$$\forall k = 1, \dots, m \left[\|\mathbf{A}\hat{\mathbf{v}}_k - \hat{\lambda}_k \hat{\mathbf{v}}_k\| \leq \varepsilon \right],$$

where $\forall k = 1, \dots, m$ $\hat{\lambda}_k$ is an approximate eigenvalue of A .

Proof. Let $P_A(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I})$ be the characteristic polynomial of A . Notice that P_A is monic of degree n and thus, by the Fundamental Theorem of Algebra, there exists a multiset of n roots $\{\lambda_1, \dots, \lambda_n\}$. Let $\hat{\mathbf{A}} = \mathbf{A}_N$ be the N -th rational approximation to \mathbf{A} . That is, for each entry a_{jk} of A , it follows that $|a_{jk} - \hat{a}_{jk}| \leq \frac{2}{N}$ where \hat{a}_{jk} is the respective entry of $\hat{\mathbf{A}}$. Since the coefficients of the characteristic polynomial are defined via the finite product and sum of the matrix entries, we can find such an N that, by using Lemma 4, the roots $\hat{\lambda}_1, \dots, \hat{\lambda}_n$ of $\hat{\mathbf{A}}$ will satisfy $\forall k = 1, \dots, n \left[|\hat{\lambda}_k - \lambda_k| \leq \varepsilon \right]$. Notice that the entries of $\hat{\mathbf{A}}$ are rational complex numbers. Thus, $\hat{\lambda}_1, \dots, \hat{\lambda}_n$ are algebraic numbers. By Lemma 4.1 from [26, p. 8], we can decide whether $\hat{\lambda}_j = \hat{\lambda}_k$ or $\hat{\lambda}_j \neq \hat{\lambda}_k$ for $j \neq k, j, k = 1, \dots, n$. Suppose that there are m distinct numbers in $\{\hat{\lambda}_1, \dots, \hat{\lambda}_n\}$, and their algebraic multiplicities are l_1, \dots, l_m . Since

$$\hat{\mathbf{A}}\hat{\mathbf{v}} = \hat{\lambda}_k \hat{\mathbf{v}}, k = 1, \dots, m$$

are systems of linear equations with coefficients being algebraic numbers, we can solve them exactly for $\hat{\mathbf{v}}$ by Gaussian elimination. Notice that these systems of linear equations are consistent, and $\hat{\mathbf{A}} - \hat{\lambda}_k I$ is singular, so there is at least a one-dimensional subspaces of vectors that satisfy the respective system of equations. Therefore, we may assume that $\|\hat{\mathbf{v}}\| = 1$. Notice that a solution $\hat{\mathbf{v}}$ is a tuple of algebraic numbers. Further, for each $\hat{\lambda}_k$, we can determine the geometric multiplicity, again, by Gaussian elimination on $\hat{\mathbf{A}} - \hat{\lambda}_k I$. If the geometric multiplicity g_k of some $\hat{\lambda}_k$ is greater than one, we can find linearly independent vectors $\hat{\mathbf{v}}_k^1 \dots \hat{\mathbf{v}}_k^{g_k}$ all satisfying $\hat{\mathbf{A}}\hat{\mathbf{v}} = \hat{\lambda}_k\hat{\mathbf{v}}$. Without loss of generality, we may assume that all geometric multiplicities are equal to one. At this point, we have a finite set of eigenpairs $(\hat{\lambda}_1, \hat{\mathbf{v}}_1), \dots, (\hat{\lambda}_m, \hat{\mathbf{v}}_m)$ of $\hat{\mathbf{A}}$. For $k = 1, \dots, m$, it follows that:

$$\begin{aligned} \|\mathbf{A}\hat{\mathbf{v}}_k - \hat{\lambda}_k\hat{\mathbf{v}}_k\| &= \|\mathbf{A}\hat{\mathbf{v}}_k + \hat{\mathbf{A}}\hat{\mathbf{v}}_k - \hat{\mathbf{A}}\hat{\mathbf{v}}_k - \hat{\lambda}_k\hat{\mathbf{v}}_k\| \\ &= \|(\mathbf{A} - \hat{\mathbf{A}})\hat{\mathbf{v}}_k + 0\| \\ &= \|(\mathbf{A} - \hat{\mathbf{A}})\hat{\mathbf{v}}_k\| \\ &\leq \|(\mathbf{A} - \hat{\mathbf{A}})\| \|\hat{\mathbf{v}}_k\| \\ &= \|(\mathbf{A} - \hat{\mathbf{A}})\| \leq n\sqrt{n}\varepsilon \end{aligned}$$

By setting N accordingly, the result follows. \square

Note that we cannot prove constructively that the approximate eigenvectors converge to the classical eigenvectors, which implies that we cannot prove existence of eigenvectors in the general case. We now prove another useful inequality. Let λ_j be such that $|\lambda_j - \hat{\lambda}_k| \leq \varepsilon$ for some k . It follows that:

$$\begin{aligned} \|\mathbf{A}\hat{\mathbf{v}}_k - \lambda_j\hat{\mathbf{v}}_k\| &= \|\mathbf{A}\hat{\mathbf{v}}_k + \hat{\mathbf{A}}\hat{\mathbf{v}}_k - \hat{\mathbf{A}}\hat{\mathbf{v}}_k + \hat{\lambda}_k\hat{\mathbf{v}}_k - \hat{\lambda}_k\hat{\mathbf{v}}_k - \lambda_j\hat{\mathbf{v}}_k\| \\ &= \|\mathbf{A}\hat{\mathbf{v}}_k - \hat{\mathbf{A}}\hat{\mathbf{v}}_k + \hat{\lambda}_k\hat{\mathbf{v}}_k - \lambda_j\hat{\mathbf{v}}_k\| \\ &\leq \|(\mathbf{A} - \hat{\mathbf{A}})\hat{\mathbf{v}}_k\| + \|(\hat{\lambda}_k - \lambda_j)\hat{\mathbf{v}}_k\| \\ &\leq \|\mathbf{A} - \hat{\mathbf{A}}\| \|\hat{\mathbf{v}}_k\| + \|\hat{\lambda}_k - \lambda_j\| \|\hat{\mathbf{v}}_k\| \\ &\leq (1 + n\sqrt{n})\varepsilon \end{aligned}$$

There exists a concept of an ε -eigenvector, introduced by [22, p. 372], in the following sense: if A is an operator on a Hilbert space H , then a vector $v \in H$ is an ε -eigenvector of A if $\|v\| = 1$ and $\|Av - (Av, v)v\| \leq \varepsilon$, where (Av, v) is the dot product. Observe that in our case, $\hat{\mathbf{A}}\hat{\mathbf{v}}_k = \hat{\lambda}_k\hat{\mathbf{v}}_k$. Transposing both sides and multiplying by $\hat{\mathbf{v}}$, we get:

$$(\hat{\mathbf{A}}\hat{\mathbf{v}}_k)^T \hat{\mathbf{v}}_k = \hat{\lambda}_k \hat{\mathbf{v}}_k^T \hat{\mathbf{v}}_k = \hat{\lambda}_k \|\hat{\mathbf{v}}_k\| = \hat{\lambda}_k.$$

Therefore, $\|\hat{\mathbf{A}}\hat{\mathbf{v}}_k - (\hat{\mathbf{A}}\hat{\mathbf{v}}_k)^T \hat{\mathbf{v}}_k\| \leq n\sqrt{n}\varepsilon$ which means that $\hat{\mathbf{v}}_k$ is an $n\sqrt{n}\varepsilon$ -eigenvector of \mathbf{A} in the sense of [22].

5 Conclusion and Discussion

In this note, we studied the computation of eigenvectors from the standpoint of constructive mathematics. The derived results allow the user to define the bound on the uncertainty ε of the eigenvectors. The computational content of the Theorem 1 is an algorithm that computes a set of unit vectors, i.e. the ε -eigenvectors. The algorithm always outputs a correct result, whereas computation in the context of the exact eigenvectors may suffer from numerical instability.

In controller design, one can guarantee robustness of computation by giving up the concept of the exact eigenvectors and focusing on ε -eigenvectors instead. When developing a controller that relies on the standard notion of exact eigenvectors (see, e.g., [1]), the user may need to manually verify that the outcome of the computation is plausible. Then, if the user suspects a numerical issue, the procedure may need to be repeated using other initial guess or another solution algorithm. In contrast, the advantage of the constructive treatment of eigenvectors and matrix decompositions lies in the fact that the respective constructions are effective, i.e. they always and automatically guarantee a correct result in terms of the given specification or predefined precision ε .

In practical applications, an essentially approximate treatment may be considered as a suitable substitute for exact eigenvectors where guaranteeing a predefined precision is essential. The derived construction meets the user specification of computation accuracy and might be useful in deriving various matrix normal forms and decompositions in an essentially approximate format. This may be especially important in applications with strict requirements for numerical stability, such as feedback control. Moreover, as formal verification methods gain popularity in control [27, 28, 29, 30], effective computations and formal proofs become more important. We therefore assume that constructive mathematics may provide a coherent answer and appropriate means to assess uncertainty of software implementation of control systems.

Acknowledgment

The authors would like to thank D. Bridges for valuable conversations on the subject of the present work.

References

- [1] M. Hovd, R. D. Braatz, and S. Skogestad. SVD controllers for h_2 , h_∞ and μ -optimal control. *Automatica*, 33(3):433 – 439, 1997.
- [2] K. C. Sou and A. Rantzer. A singular value decomposition based closed loop stability preserving controller reduction method. In *American Control Conference (ACC), 2010*, pages 1079–1084. IEEE, 2010.

- [3] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems*, volume 6. SIAM, 2005.
- [4] O. J. Rojas, G. C. Goodwin, M. M. Serón, and A. Feuer. An SVD based strategy for receding horizon control of input constrained linear systems. *International Journal of Robust and Nonlinear Control*, 14(13-14):1207–1226, 2004.
- [5] Y. Zhang, G. Dai, H. Zhang, and Q. Li. A SVD-based extended kalman filter and applications to aircraft flight state and parameter estimation. In *American Control Conference, 1994*, volume 2, pages 1809–1813 vol.2, June 1994.
- [6] C. E. Rohrs, D. Schultz, and J. Melsa. *Linear Control Systems*. McGraw-Hill Higher Education, 1st edition, 1992.
- [7] D. M. Walker, G. Froyland, K. Judd, and A. I. Mees. Controllers for nonlinear systems using normal forms. *International Journal of Bifurcation and Chaos*, 13(02):459–465, 2003.
- [8] D. C. Lay. *Linear Algebra and Its Applications*. Addison-Wesley, 2012.
- [9] A. M. Turing. On computable numbers, with an application to the entscheidungsproblem. *J. of Math*, 58(345-363):5, 1936.
- [10] A. Neumaier. Solving ill-conditioned and singular linear systems: A tutorial on regularization. *SIAM review*, 40(3):636–666, 1998.
- [11] Sh. Chandrasekaran and I. Ipsen. Perturbation theory for the solution of systems of linear equations. Technical report, DTIC Document, 1991.
- [12] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola. *Numerical Methods for the Solution of Ill-posed Problems*, volume 328. Springer Science & Business Media, 1995.
- [13] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Mathematical Modeling and Computation. Society for Industrial and Applied Mathematics, 1998.
- [14] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Mathematics and Its Applications. Springer Netherlands, 1996.
- [15] D. Trefethen, L. Bau. *Numerical Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, 1997.
- [16] David Hecker and Deborah Lurie. Using least-squares to find an approximate eigenvector. *ELA. The Electronic Journal of Linear Algebra [electronic only]*, 16:99–110, 2007.
- [17] K. Weihrauch. *Computable Analysis: an Introduction*. Springer Science & Business Media, 2012.

- [18] V. Brattka and M. Ziegler. Computability of linear equations. In *Foundations of Information Technology in the Era of Network and Mobile Computing*, pages 95–106. Springer, 2002.
- [19] M. Ziegler and V. Brattka. Computing the dimension of linear subspaces. In *SOFSEM 2000: Theory and Practice of Informatics*, pages 450–458. Springer, 2000.
- [20] R. Mines, F. Richman, and W. Ruitenburg. *A Course in Constructive Algebra*. Springer Science & Business Media, 1988.
- [21] F. Ye. *Strict Finitism and the Logic of Mathematical Applications*. Springer, 2011.
- [22] E. Bishop and D. Bridges. *Constructive Analysis*, volume 279. Springer Science & Business Media, 1985.
- [23] D.S. Bridges and L.S. Vita. *Techniques of Constructive Analysis*. Universitext. Springer New York, 2007.
- [24] D. Bridges and F. Richman. *Varieties of Constructive Mathematics*, volume 97. Cambridge University Press, 1987.
- [25] A. Beardon. *Complex Analysis: the Winding Number Principle in Analysis and Topology*. Wiley, New York, 1979.
- [26] M. J. Beeson. *Foundations of Constructive Mathematics: Metamathematical Studies*, volume 6. Springer Science & Business Media, 1980.
- [27] C. Bernardeschi and A. Domenici. Verifying safety properties of a nonlinear control by interactive theorem proving with the prototype verification system. *Information Processing Letters*, 116(6):409 – 415, 2016.
- [28] D Araiza-Illan, K. Eder, and A. Richards. Formal verification of control systems’ properties with theorem proving. In *Control (CONTROL), 2014 UKACC International Conference on*, pages 244–249, July 2014.
- [29] W. Damm, A. Mikschl, J. Oehlerking, E.-R. Olderog, J. Pang, A. Platzer, M. Segelken, and B. Wirtz. Automating verification of cooperation, control, and design in traffic applications. In *Formal Methods and Hybrid Real-time Systems. Volume 4700 of Lecture Notes in Computer Science*, pages 115–169. Springer, 2007.
- [30] S. Gao. Descriptive control theory: A proposal. *CoRR*, abs/1409.3560, 2014.